SEQUENCE LISTING

<110> Palmer, Michelle A.J.
      Gee, Melissa
      Tillotson, Bonnie
      Chang, Xiao-Jia

<120> Receptor Function Assay for G-Protein Coupled Receptors and Orphan
      Receptors by Reporter Enzyme Mutant Complementation

<130> 4085-226-27

<140> US 09/654,499
<141> 2000-09-01

<150> US 60/180,669
<151> 2000-02-07

<160> 6

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 6700
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence for pICAST ALC

<221> CDS
<222> (1457)...(4486)

<400> 1
ctgcagcctg aatatgggcc aaacaggata tctgtggtaa gcagttcctg ccccggctca 60
gggccaagaa cagatggaac agctgaatat gggccaaaca ggatatctgt ggtaagcagt 120
tcctgccccg gctcagggcc aagaacagat ggtccccaga tgcggtccag ccctcagcag 180
tttctagaga accatcagat gtttccaggg tgccccaagg acctgaaatg accctgtgcc 240
ttatttgaac taaccaatca gttcgcttct cgcttctgtt cgcgcgcttc tgctccccga 300
gctcaataaa agagcccaca acccctcact cggggcgcca gtcctccgat tgactgagtc 360
gcccgggtac ccgtgtatcc aataaaccct cttgcagttg catccgactt gtggtctcgc 420
tgttccttgg gagggtctcc tctgagtgat tgactacccg tcagcggggg tctttcattt 480
gggggctcgt ccgggatcgg gagacccctg cccagggacc accgacccac caccgggagg 540
caagctggcc agcaacttat ctgtgtctgt ccgattgtct agtgtctatg actgattta 600
tgcgcctgcg tcggtactag ttagctaact agctctgtat ctggcggacc cgtggtggaa 660
ctgacgagtt ctgaacaccc ggccgcaacc ctgggagacg tcccagggac tttggggggcc 720
gtttttgtgg cccgacctga ggaagggagt cgatgtggaa tccgaccccg tcaggatatg 780
tggttctggt aggagacgag aacctaaaac agttcccgcc tccgtctgaa tttttgcttt 840
cggtttggaa ccgaagccgc gcgtcttgtc tgctgcagca tcgttctgtg ttgtctctgt 900
ctgactgtgt ttctgtattt gtctgaaaat tagggccaga ctgttaccac tcccttaagt 960
ttgaccttag gtaactggaa agatgtcgag cggctcgctc acaaccagtc ggtagatgtc 1020
aagaagagac gttgggttac cttctgctct gcagaatggc caacctttaa cgtcggatgg 1080
ccgcgagacg gcacctttaa ccgagacctc atcacccagg ttaagatcaa ggtctttttca 1140
cctggccccgc atggacaccc agaccaggtc ccctacatcg tgacctggga agccttggct 1200
tttgacccccc ctccctgggt caagcccttt gtacaccccta agcctccgcc tcctcttcct 1260
ccatccgccc cgtctctccc ccttgaacct cctcgttcga ccccgcctcg atcctccctt 1320
tatccagccc tcactccttc tctaggcgcc ggccgctcta gcccattaat acgactcact 1380
atagggcgat tcgaatcagg ccttggcgcg ccggatcctt aattaagcgc aattgggagg 1440

```
tggcggtagc ctcgag atg ggc gtg att acg gat tca ctg gcc gtc gtg gcc 1492
                   Met Gly Val Ile Thr Asp Ser Leu Ala Val Val Ala
                    1               5                   10

cgc acc gat cgc cct tcc caa cag tta cgc agc ctg aat ggc gaa tgg    1540
Arg Thr Asp Arg Pro Ser Gln Gln Leu Arg Ser Leu Asn Gly Glu Trp
        15                  20                  25

cgc ttt gcc tgg ttt ccg gca cca gaa gcg gtg ccg gaa agc tgg ctg    1588
Arg Phe Ala Trp Phe Pro Ala Pro Glu Ala Val Pro Glu Ser Trp Leu
        30                  35                  40

gag tgc gat ctt cct gag gcc gat act gtc gtc gtc ccc tca aac tgg    1636
Glu Cys Asp Leu Pro Glu Ala Asp Thr Val Val Val Pro Ser Asn Trp
45                  50                  55                  60

cag atg cac ggt tac gat gcg ccc atc tac acc aac gtg acc tat ccc    1684
Gln Met His Gly Tyr Asp Ala Pro Ile Tyr Thr Asn Val Thr Tyr Pro
                65                  70                  75

att acg gtc aat ccg ccg ttt gtt ccc acg gag aat ccg acg ggt tgt    1732
Ile Thr Val Asn Pro Pro Phe Val Pro Thr Glu Asn Pro Thr Gly Cys
                80                  85                  90

tac tcg ctc aca ttt aat gtt gat gaa agc tgg cta cag gaa ggc cag    1780
Tyr Ser Leu Thr Phe Asn Val Asp Glu Ser Trp Leu Gln Glu Gly Gln
                95                  100                 105

acg cga att att ttt gat ggc gtt aac tcg gcg ttt cat ctg tgg tgc    1828
Thr Arg Ile Ile Phe Asp Gly Val Asn Ser Ala Phe His Leu Trp Cys
        110                 115                 120

aac ggg cgc tgg gtc ggt tac ggc cag gac agt cgt ttg ccg tct gaa    1876
Asn Gly Arg Trp Val Gly Tyr Gly Gln Asp Ser Arg Leu Pro Ser Glu
125                 130                 135                 140

ttt gac ctg agc gca ttt tta cgc gcc gga gaa aac cgc ctc gcg gtg    1924
Phe Asp Leu Ser Ala Phe Leu Arg Ala Gly Glu Asn Arg Leu Ala Val
                145                 150                 155

atg gtg ctg cgc tgg agt gac ggc agt tat ctg gaa gat cag gat atg    1972
Met Val Leu Arg Trp Ser Asp Gly Ser Tyr Leu Glu Asp Gln Asp Met
                160                 165                 170

tgg cgg atg agc ggc att ttc cgt gac gtc tcg ttg ctg cat aaa ccg    2020
Trp Arg Met Ser Gly Ile Phe Arg Asp Val Ser Leu Leu His Lys Pro
        175                 180                 185

act aca caa atc agc gat ttc cat gtt gcc act cgc ttt aat gat gat    2068
Thr Thr Gln Ile Ser Asp Phe His Val Ala Thr Arg Phe Asn Asp Asp
        190                 195                 200

ttc agc cgc gct gta ctg gag gct gaa gtt cag atg tgc ggc gag ttg    2116
Phe Ser Arg Ala Val Leu Glu Ala Glu Val Gln Met Cys Gly Glu Leu
205                 210                 215                 220

cgt gac tac cta cgg gta aca gtt tct tta tgg cag ggt gaa acg cag    2164
Arg Asp Tyr Leu Arg Val Thr Val Ser Leu Trp Gln Gly Glu Thr Gln
                225                 230                 235
```

```
gtc gcc agc ggc acc gcg cct ttc ggc ggt gaa att atc gat gag cgt    2212
Val Ala Ser Gly Thr Ala Pro Phe Gly Gly Glu Ile Ile Asp Glu Arg
            240                 245                 250

ggt ggt tat gcc gat cgc gtc aca cta cgt ctg aac gtc gaa aac ccg    2260
Gly Gly Tyr Ala Asp Arg Val Thr Leu Arg Leu Asn Val Glu Asn Pro
            255                 260                 265

aaa ctg tgg agc gcc gaa atc ccg aat ctc tat cgt gcg gtg gtt gaa    2308
Lys Leu Trp Ser Ala Glu Ile Pro Asn Leu Tyr Arg Ala Val Val Glu
        270                 275                 280

ctg cac acc gcc gac ggc acg ctg att gaa gca gaa gcc tgc gat gtc    2356
Leu His Thr Ala Asp Gly Thr Leu Ile Glu Ala Glu Ala Cys Asp Val
285                 290                 295                 300

ggt ttc cgc gag gtg cgg att gaa aat ggt ctg ctg ctg ctg aac ggc    2404
Gly Phe Arg Glu Val Arg Ile Glu Asn Gly Leu Leu Leu Leu Asn Gly
                305                 310                 315

aag ccg ttg ctg att cga ggc gtt aac cgt cac gag cat cat cct ctg    2452
Lys Pro Leu Leu Ile Arg Gly Val Asn Arg His Glu His His Pro Leu
            320                 325                 330

cat ggt cag gtc atg gat gag cag acg atg gtg cag gat atc ctg ctg    2500
His Gly Gln Val Met Asp Glu Gln Thr Met Val Gln Asp Ile Leu Leu
            335                 340                 345

atg aag cag aac aac ttt aac gcc gtg cgc tgt tcg cat tat ccg aac    2548
Met Lys Gln Asn Asn Phe Asn Ala Val Arg Cys Ser His Tyr Pro Asn
        350                 355                 360

cat ccg ctg tgg tac acg ctg tgc gac cgc tac ggc ctg tat gtg gtg    2596
His Pro Leu Trp Tyr Thr Leu Cys Asp Arg Tyr Gly Leu Tyr Val Val
365                 370                 375                 380

gat gaa gcc aat att gaa acc cac ggc atg gtg cca atg aat cgt ctg    2644
Asp Glu Ala Asn Ile Glu Thr His Gly Met Val Pro Met Asn Arg Leu
                385                 390                 395

acc gat gat ccg cgc tgg cta ccg gcg atg agc gaa cgc gta acg cga    2692
Thr Asp Asp Pro Arg Trp Leu Pro Ala Met Ser Glu Arg Val Thr Arg
            400                 405                 410

atg gtg cag cgc gat cgt aat cac ccg agt gtg atc atc tgg tcg ctg    2740
Met Val Gln Arg Asp Arg Asn His Pro Ser Val Ile Ile Trp Ser Leu
            415                 420                 425

ggg aat gaa tca ggc cac ggc gct aat cac gac gcg ctg tat cgc tgg    2788
Gly Asn Glu Ser Gly His Gly Ala Asn His Asp Ala Leu Tyr Arg Trp
        430                 435                 440

atc aaa tct gtc gat cct tcc cgc ccg gtg cag tat gaa ggc ggc gga    2836
Ile Lys Ser Val Asp Pro Ser Arg Pro Val Gln Tyr Glu Gly Gly Gly
445                 450                 455                 460

gcc gac acc acg gcc acc gat att att tgc ccg atg tac gcg cgc gtg    2884
Ala Asp Thr Thr Ala Thr Asp Ile Ile Cys Pro Met Tyr Ala Arg Val
            465                 470                 475
```

```
gat gaa gac cag ccc ttc ccg gct gtg ccg aaa tgg tcc atc aaa aaa    2932
Asp Glu Asp Gln Pro Phe Pro Ala Val Pro Lys Trp Ser Ile Lys Lys
            480                 485                 490

tgg ctt tcg cta cct gga gag acg cgc ccg ctg atc ctt tgc gaa tac    2980
Trp Leu Ser Leu Pro Gly Glu Thr Arg Pro Leu Ile Leu Cys Glu Tyr
            495                 500                 505

gcc cac gcg atg ggt aac agt ctt ggc ggt ttc gct aaa tac tgg cag    3028
Ala His Ala Met Gly Asn Ser Leu Gly Gly Phe Ala Lys Tyr Trp Gln
    510                 515                 520

gcg ttt cgt cag tat ccc cgt tta cag ggc ggc ttc gtc tgg gac tgg    3076
Ala Phe Arg Gln Tyr Pro Arg Leu Gln Gly Gly Phe Val Trp Asp Trp
525                 530                 535                 540

gtg gat cag tcg ctg att aaa tat gat gaa aac ggc aac ccg tgg tcg    3124
Val Asp Gln Ser Leu Ile Lys Tyr Asp Glu Asn Gly Asn Pro Trp Ser
            545                 550                 555

gct tac ggc ggt gat ttt ggc gat acg ccg aac gat cgc cag ttc tgt    3172
Ala Tyr Gly Gly Asp Phe Gly Asp Thr Pro Asn Asp Arg Gln Phe Cys
            560                 565                 570

atg aac ggt ctg gtc ttt gcc gac cgc acg ccg cat cca gcg ctg acg    3220
Met Asn Gly Leu Val Phe Ala Asp Arg Thr Pro His Pro Ala Leu Thr
            575                 580                 585

gaa gca aaa cac cag cag cag ttt ttc cag ttc cgt tta tcc ggg caa    3268
Glu Ala Lys His Gln Gln Gln Phe Phe Gln Phe Arg Leu Ser Gly Gln
    590                 595                 600

acc atc gaa gtg acc agc gaa tac ctg ttc cgt cat agc gat aac gag    3316
Thr Ile Glu Val Thr Ser Glu Tyr Leu Phe Arg His Ser Asp Asn Glu
605                 610                 615                 620

ctc ctg cac tgg atg gtg gcg ctg gat ggt aag ccg ctg gca agc ggt    3364
Leu Leu His Trp Met Val Ala Leu Asp Gly Lys Pro Leu Ala Ser Gly
                625                 630                 635

gaa gtg cct ctg gat gtc gct cca caa ggt aaa cag ttg att gaa ctg    3412
Glu Val Pro Leu Asp Val Ala Pro Gln Gly Lys Gln Leu Ile Glu Leu
            640                 645                 650

cct gaa cta ccg cag ccg gag agc gcc ggg caa ctc tgg ctc aca gta    3460
Pro Glu Leu Pro Gln Pro Glu Ser Ala Gly Gln Leu Trp Leu Thr Val
            655                 660                 665

cgc gta gtg caa ccg aac gcg acc gca tgg tca gaa gcc ggg cac atc    3508
Arg Val Val Gln Pro Asn Ala Thr Ala Trp Ser Glu Ala Gly His Ile
    670                 675                 680

agc gcc tgg cag cag tgg cgt ctg gcg gaa aac ctc agt gtg acg ctc    3556
Ser Ala Trp Gln Gln Trp Arg Leu Ala Glu Asn Leu Ser Val Thr Leu
685                 690                 695                 700

ccc gcc gcg tcc cac gcc atc ccg cat ctg acc acc agc gaa atg gat    3604
Pro Ala Ala Ser His Ala Ile Pro His Leu Thr Thr Ser Glu Met Asp
            705                 710                 715
```

```
ttt tgc atc gag ctg ggt aat aag cgt tgg caa ttt aac cgc cag tca    3652
Phe Cys Ile Glu Leu Gly Asn Lys Arg Trp Gln Phe Asn Arg Gln Ser
            720             725             730

ggc ttt ctt tca cag atg tgg att ggc gat aaa aaa caa ctg ctg acg    3700
Gly Phe Leu Ser Gln Met Trp Ile Gly Asp Lys Lys Gln Leu Leu Thr
            735             740             745

ccg ctg cgc gat cag ttc acc cgt gca ccg ctg gat aac gac att ggc    3748
Pro Leu Arg Asp Gln Phe Thr Arg Ala Pro Leu Asp Asn Asp Ile Gly
        750             755             760

gta agt gaa gcg acc cgc att gac cct aac gcc tgg gtc gaa cgc tgg    3796
Val Ser Glu Ala Thr Arg Ile Asp Pro Asn Ala Trp Val Glu Arg Trp
765             770             775             780

aag gcg gcg ggc cat tac cag gcc gaa gca gcg ttg ttg cag tgc acg    3844
Lys Ala Ala Gly His Tyr Gln Ala Glu Ala Ala Leu Leu Gln Cys Thr
            785             790             795

gca gat aca ctt gct gat gcg gtg ctg att acg acc gct cac gcg tgg    3892
Ala Asp Thr Leu Ala Asp Ala Val Leu Ile Thr Thr Ala His Ala Trp
            800             805             810

cag cat cag ggg aaa acc tta ttt atc agc cgg aaa acc tac cgg att    3940
Gln His Gln Gly Lys Thr Leu Phe Ile Ser Arg Lys Thr Tyr Arg Ile
            815             820             825

gat ggt agt ggt caa atg gcg att acc gtt gat gtt gaa gtg gcg agc    3988
Asp Gly Ser Gly Gln Met Ala Ile Thr Val Asp Val Glu Val Ala Ser
        830             835             840

gat aca ccg cat ccg gcg cgg att ggc ctg aac tgc cag ctg gcg cag    4036
Asp Thr Pro His Pro Ala Arg Ile Gly Leu Asn Cys Gln Leu Ala Gln
845             850             855             860

gta gca gag cgg gta aac tgg ctc gga tta ggg ccg caa gaa aac tat    4084
Val Ala Glu Arg Val Asn Trp Leu Gly Leu Gly Pro Gln Glu Asn Tyr
            865             870             875

ccc gac cgc ctt act gcc gcc tgt ttt gac cgc tgg gat ctg cca ttg    4132
Pro Asp Arg Leu Thr Ala Ala Cys Phe Asp Arg Trp Asp Leu Pro Leu
            880             885             890

tca gac atg tat acc ccg tac gtc ttc ccg agc gaa aac ggt ctg cgc    4180
Ser Asp Met Tyr Thr Pro Tyr Val Phe Pro Ser Glu Asn Gly Leu Arg
            895             900             905

tgc ggg acg cgc gaa ttg aat tat ggc cca cac cag tgg cgc ggc gac    4228
Cys Gly Thr Arg Glu Leu Asn Tyr Gly Pro His Gln Trp Arg Gly Asp
            910             915             920

ttc cag ttc aac atc agc cgc tac agt caa cag caa ctg atg gaa acc    4276
Phe Gln Phe Asn Ile Ser Arg Tyr Ser Gln Gln Gln Leu Met Glu Thr
925             930             935             940

agc cat cgc cat ctg ctg cac gcg gaa gaa ggc aca tgg ctg aat atc    4324
Ser His Arg His Leu Leu His Ala Glu Glu Gly Thr Trp Leu Asn Ile
            945             950             955
```

```
gac ggt ttc cat atg ggg att ggt ggc gac gac tcc tgg agc ccg tca    4372
Asp Gly Phe His Met Gly Ile Gly Gly Asp Asp Ser Trp Ser Pro Ser
            960                 965                 970

gta tcg gcg gaa ttc cag ctg agc gcc ggt cgc tac cat tac cag ttg    4420
Val Ser Ala Glu Phe Gln Leu Ser Ala Gly Arg Tyr His Tyr Gln Leu
        975                 980                 985

gtc tgg tgt caa aaa aga tct gac tat aaa gat gag gac ctc gac cat    4468
Val Trp Cys Gln Lys Arg Ser Asp Tyr Lys Asp Glu Asp Leu Asp His
     990                 995                 1000

cat cat cat cat cac cgg taataatagg tagataagtg actgattaga           4516
His His His His His Arg
1005                1010

tgcattgatc cctcgaccaa ttccggttat tttccaccat attgccgtct tttggcaatg 4576
tgagggcccg aaacctggc cctgtcttct tgacgagcat tcctaggggt ctttcccctc 4636
tcgccaaagg aatgcaaggt ctgttgaatg tcgtgaagga agcagttcct ctggaagctt 4696
cttgaagaca acaacgtct gtagcgaccc tttgcaggca gcggaacccc ccacctggcg 4756
acaggtgcct ctgcggccaa aagccacgtg tataagatac acctgcaaag gcggcacaac 4816
cccagtgcca cgttgtgagt tggatagttg tggaaagagt caaatggctc tcctcaagcg 4876
tattcaacaa ggggctgaag gatgcccaga aggtacccca ttgtatggga tctgatctgg 4936
ggcctcggtg cacatgcttt acatgtgttt agtcgaggtt aaaaaacgtc taggcccccc 4996
gaaccacggg gacgtggttt tcctttgaaa aacacgatga taataccatg attgaacaag 5056
atggattgca cgcaggttct ccggccgctt gggtggagag ctattcgcc tatgactggg 5116
cacaacagac aatcggctgc tctgatgccg ccgtgttccg gctgtcagcg caggggcgcc 5176
cggttctttt tgtcaagacc gacctgtccg gtgccctgaa tgaactgcag gacgaggcag 5236
cgcggctatc gtggctggcc acgacgggcg ttccttgccg agctgtgctc gacgttgtca 5296
ctgaagcggg aagggactgg ctgctattgg gcgaagtgcc ggggcaggat ctcctgtcat 5356
ctcaccttgc tcctgccgag aaagtatcca tcatggctga tgcaatgcgg cggctgcata 5416
cgcttgatcc ggctacctgc ccattcgacc accaagcgaa acatcgcatc gagcgagcac 5476
gtactcggat ggaagccggt cttgtcgatc aggatgatct ggacgaagag catcagggc 5536
tcgcgccagc cgaactgttc gccaggctca aggcgcgcat gcccgacggc gaggatctcg 5596
tcgtgaccca tggcgatgcc tgcttgccga atatcatggt ggaaaatggc cgcttttctg 5656
gattcatcga ctgtggccgg ctgggtgtgg cggaccgcta tcaggacata gcgttggcta 5716
cccgtgatat tgctgaagag cttggcggcg aatgggctga ccgcttcctc gtgctttacg 5776
gtatcgccgc tcccgattcg cagcgcatcg ccttctatcg ccttcttgac gagttcttct 5836
gagcgggact ctggggttcg catcgataaa ataaagatt ttatttagtc tccagaaaaa 5896
ggggggaatg aaagaccca cctgtaggtt tggcaagcta gcttaagtaa cgccattttg 5956
caaggcatgg aaaaatacat aactgagaat agagaagttc agatcaaggt caggaacaga 6016
tggaacagct gaatatgggc caaacaggat atctgtggta agcagttcct gccccggctc 6076
agggccaaga acagatggaa cagctgaata tgggccaaac aggatatctg tggtaagcag 6136
ttcctgcccc ggctcagggc caagaacaga tggtccccag atgcggtcca gccctcagca 6196
gtttctagag aaccatcaga tgtttccagg gtgccccaag acctgaaat gaccctgtgc 6256
cttatttgaa ctaaccaatc agttcgcttc tcgcttctgt tcgcgcgctt ctgctccccg 6316
agctcaataa aagagcccac aacccctcac tcggggcgcc agtcctccga ttgactgagt 6376
cgcccgggta cccgtgtatc caataaaccc tcttgcagtt gcatccgact tgtggtctcg 6436
ctgttccttg ggagggtctc ctctgagtga ttgactaccc gtcagcgggg gtctttcatt 6496
catgcagcat gtatcaaaat taatttggtt tttttctta agtatttaca ttaaatggcc 6556
atagttgcat taatgaatcg gccaacgcgc ggggagaggc ggtttgcgta ttggcgctct 6616
tccgcttcct cgctcactga ctcgctgcgc tcggtcgttc ggctgcggcg agcggtatca 6676
gctcactcaa aggcggtaat acgg                                         6700
```

<220>
<223> Amino acid sequence for pICAST ALC

<400> 2

```
Met Gly Val Ile Thr Asp Ser Leu Ala Val Val Ala Arg Thr Asp Arg
 1               5                  10                  15
Pro Ser Gln Gln Leu Arg Ser Leu Asn Gly Glu Trp Arg Phe Ala Trp
                20                  25                  30
Phe Pro Ala Pro Glu Ala Val Pro Glu Ser Trp Leu Glu Cys Asp Leu
        35                  40                  45
Pro Glu Ala Asp Thr Val Val Val Pro Ser Asn Trp Gln Met His Gly
    50                  55                  60
Tyr Asp Ala Pro Ile Tyr Thr Asn Val Thr Tyr Pro Ile Thr Val Asn
65                  70                  75                  80
Pro Pro Phe Val Pro Thr Glu Asn Pro Thr Gly Cys Tyr Ser Leu Thr
                85                  90                  95
Phe Asn Val Asp Glu Ser Trp Leu Gln Glu Gly Gln Thr Arg Ile Ile
                100                 105                 110
Phe Asp Gly Val Asn Ser Ala Phe His Leu Trp Cys Asn Gly Arg Trp
        115                 120                 125
Val Gly Tyr Gly Gln Asp Ser Arg Leu Pro Ser Glu Phe Asp Leu Ser
    130                 135                 140
Ala Phe Leu Arg Ala Gly Glu Asn Arg Leu Ala Val Met Val Leu Arg
145                 150                 155                 160
Trp Ser Asp Gly Ser Tyr Leu Glu Asp Gln Asp Met Trp Arg Met Ser
                165                 170                 175
Gly Ile Phe Arg Asp Val Ser Leu Leu His Lys Pro Thr Thr Gln Ile
                180                 185                 190
Ser Asp Phe His Val Ala Thr Arg Phe Asn Asp Asp Phe Ser Arg Ala
        195                 200                 205
Val Leu Glu Ala Glu Val Gln Met Cys Gly Glu Leu Arg Asp Tyr Leu
    210                 215                 220
Arg Val Thr Val Ser Leu Trp Gln Gly Glu Thr Gln Val Ala Ser Gly
225                 230                 235                 240
Thr Ala Pro Phe Gly Gly Glu Ile Ile Asp Glu Arg Gly Gly Tyr Ala
                245                 250                 255
Asp Arg Val Thr Leu Arg Leu Asn Val Glu Asn Pro Lys Leu Trp Ser
                260                 265                 270
Ala Glu Ile Pro Asn Leu Tyr Arg Ala Val Val Glu Leu His Thr Ala
                275                 280                 285
Asp Gly Thr Leu Ile Glu Ala Glu Ala Cys Asp Val Gly Phe Arg Glu
        290                 295                 300
Val Arg Ile Glu Asn Gly Leu Leu Leu Leu Asn Gly Lys Pro Leu Leu
305                 310                 315                 320
Ile Arg Gly Val Asn Arg His Glu His His Pro Leu His Gly Gln Val
                325                 330                 335
Met Asp Glu Gln Thr Met Val Gln Asp Ile Leu Leu Met Lys Gln Asn
        340                 345                 350
Asn Phe Asn Ala Val Arg Cys Ser His Tyr Pro Asn His Pro Leu Trp
        355                 360                 365
Tyr Thr Leu Cys Asp Arg Tyr Gly Leu Tyr Val Val Asp Glu Ala Asn
370                 375                 380
Ile Glu Thr His Gly Met Val Pro Met Asn Arg Leu Thr Asp Asp Pro
385                 390                 395                 400
Arg Trp Leu Pro Ala Met Ser Glu Arg Val Thr Arg Met Val Gln Arg
                405                 410                 415
Asp Arg Asn His Pro Ser Val Ile Ile Trp Ser Leu Gly Asn Glu Ser
                420                 425                 430
Gly His Gly Ala Asn His Asp Ala Leu Tyr Arg Trp Ile Lys Ser Val
        435                 440                 445
Asp Pro Ser Arg Pro Val Gln Tyr Glu Gly Gly Gly Ala Asp Thr Thr
    450                 455                 460
```

```
Ala Thr Asp Ile Ile Cys Pro Met Tyr Ala Arg Val Asp Glu Asp Gln
465                 470             475                 480
Pro Phe Pro Ala Val Pro Lys Trp Ser Ile Lys Lys Trp Leu Ser Leu
            485             490                 495
Pro Gly Glu Thr Arg Pro Leu Ile Leu Cys Glu Tyr Ala His Ala Met
        500             505             510
Gly Asn Ser Leu Gly Gly Phe Ala Lys Tyr Trp Gln Ala Phe Arg Gln
        515             520             525
Tyr Pro Arg Leu Gln Gly Gly Phe Val Trp Asp Trp Val Asp Gln Ser
    530             535             540
Leu Ile Lys Tyr Asp Glu Asn Gly Asn Pro Trp Ser Ala Tyr Gly Gly
545             550             555                 560
Asp Phe Gly Asp Thr Pro Asn Asp Arg Gln Phe Cys Met Asn Gly Leu
            565             570             575
Val Phe Ala Asp Arg Thr Pro His Pro Ala Leu Thr Glu Ala Lys His
        580             585             590
Gln Gln Gln Phe Phe Gln Phe Arg Leu Ser Gly Gln Thr Ile Glu Val
        595             600             605
Thr Ser Glu Tyr Leu Phe Arg His Ser Asp Asn Glu Leu Leu His Trp
    610             615             620
Met Val Ala Leu Asp Gly Lys Pro Leu Ala Ser Gly Glu Val Pro Leu
625             630             635                 640
Asp Val Ala Pro Gln Gly Lys Gln Leu Ile Glu Leu Pro Glu Leu Pro
            645             650             655
Gln Pro Glu Ser Ala Gly Gln Leu Trp Leu Thr Val Arg Val Val Gln
        660             665             670
Pro Asn Ala Thr Ala Trp Ser Glu Ala Gly His Ile Ser Ala Trp Gln
        675             680             685
Gln Trp Arg Leu Ala Glu Asn Leu Ser Val Thr Leu Pro Ala Ala Ser
    690             695             700
His Ala Ile Pro His Leu Thr Thr Ser Glu Met Asp Phe Cys Ile Glu
705             710             715                 720
Leu Gly Asn Lys Arg Trp Gln Phe Asn Arg Gln Ser Gly Phe Leu Ser
            725             730             735
Gln Met Trp Ile Gly Asp Lys Lys Gln Leu Leu Thr Pro Leu Arg Asp
        740             745             750
Gln Phe Thr Arg Ala Pro Leu Asp Asn Asp Ile Gly Val Ser Glu Ala
    755             760             765
Thr Arg Ile Asp Pro Asn Ala Trp Val Glu Arg Trp Lys Ala Ala Gly
770             775             780
His Tyr Gln Ala Glu Ala Ala Leu Leu Gln Cys Thr Ala Asp Thr Leu
785             790             795                 800
Ala Asp Ala Val Leu Ile Thr Thr Ala His Ala Trp Gln His Gln Gly
            805             810             815
Lys Thr Leu Phe Ile Ser Arg Lys Thr Tyr Arg Ile Asp Gly Ser Gly
        820             825             830
Gln Met Ala Ile Thr Val Asp Val Glu Val Ala Ser Asp Thr Pro His
        835             840             845
Pro Ala Arg Ile Gly Leu Asn Cys Gln Leu Ala Gln Val Ala Glu Arg
    850             855             860
Val Asn Trp Leu Gly Leu Gly Pro Gln Glu Asn Tyr Pro Asp Arg Leu
865             870             875                 880
Thr Ala Ala Cys Phe Asp Arg Trp Asp Leu Pro Leu Ser Asp Met Tyr
            885             890             895
Thr Pro Tyr Val Phe Pro Ser Glu Asn Gly Leu Arg Cys Gly Thr Arg
        900             905             910
Glu Leu Asn Tyr Gly Pro His Gln Trp Arg Gly Asp Phe Gln Phe Asn
        915             920             925
Ile Ser Arg Tyr Ser Gln Gln Gln Leu Met Glu Thr Ser His Arg His
    930             935             940
Leu Leu His Ala Glu Glu Gly Thr Trp Leu Asn Ile Asp Gly Phe His
945             950             955                 960
```

```
Met Gly Ile Gly Gly Asp Asp Ser Trp Ser Pro Ser Val Ser Ala Glu
              965                     970                 975
Phe Gln Leu Ser Ala Gly Arg Tyr His Tyr Gln Leu Val Trp Cys Gln
              980                     985                 990
Lys Arg Ser Asp Tyr Lys Asp Glu Asp Leu Asp His His His His His
          995                 1000                1005
His Arg
    1010
```

```
<210> 3
<211> 8518
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence for pICAST ALN

<400> 3
ctgcagcctg aatatgggcc aaacaggata tctgtggtaa gcagttcctg ccccggctca 60
gggccaagaa cagatggaac agctgaatat gggccaaaca ggatatctgt ggtaagcagt 120
tcctgccccg gctcagggcc aagaacagat ggtccccaga tgcggtccag ccctcagcag 180
tttctagaga accatcagat gtttccaggg tgccccaagg acctgaaatg accctgtgcc 240
ttatttgaac taaccaatca gttcgcttct cgcttctgtt cgcgcgcttc tgctccccga 300
gctcaataaa agagcccaca acccctcact cggggcgcca gtcctccgat tgactgagtc 360
gcccgggtac ccgtgtatcc aataaaccct cttgcagttg catccgactt gtggtctcgc 420
tgttccttgg gagggtctcc tctgagtgat tgactacccg tcagcggggg tctttcattt 480
gggggctcgt ccgggatcgg agacccctg cccagggacc accgacccac caccgggagg 540
caagctggcc agcaactat ctgtgtctgt ccgattgtct agtgtctatg actgattta 600
tgcgcctgcg tcggtactag ttagctaact agctctgtat ctggcggacc cgtggtggaa 660
ctgacgagtt ctgaacaccc ggccgcaacc ctgggagacg tcccagggac tttggggggcc 720
gttttttgtgg cccgacctga ggaagggagt cgatgtggaa tccgaccccg tcaggatatg 780
tggttctggt aggagacgaa aacctaaaac agttcccgcc tccgtctgaa tttttgcttt 840
cggtttggaa ccgaagccgc gcgtcttgtc tgctgcagca tcgttctgtg ttgtctctgt 900
ctgactgtgt ttctgtattt gtctgaaaat tagggccaga ctgttaccac tcccttaagt 960
ttgaccttag gtaactggaa agatgtcgag cggctcgctc acaaccagtc ggtagatgtc 1020
aagaagagac gttgggttac cttctgctct gcagaatggc caacctttaa cgtcggatgg 1080
ccgcgagacg gcacctttaa ccgagacctc atcacccagg ttaagatcaa ggtctttttca 1140
cctggcccgc atggacaccc agaccaggtc ccctacatcg tgacctggga agccttggct 1200
tttgacccc ctccctgggt caagccctt gtacaccta agcctccgcc tcctcttcct 1260
ccatccgccc cgtctctccc ccttgaacct cctcgttcga ccccgcctcg atcctccctt 1320
tatccagccc tcactccttc tctaggcgcc ggccgctcta gcccattaat acgactcact 1380
ataggggcgat tcgaacacca tgcaccatca tcatcatcac gtcgactata aagatgagga 1440
cctcgagatg ggcgtgatta cggattcact ggccgtcgtg gcccgcaccg atcgcccttc 1500
ccaacagtta cgcagcctga atggcgaatg cgctttgcc tggtttccgg caccagaagc 1560
ggtgccggaa agctggctgg agtgcgatct tcctgaggcc gatactgtcg tcgtcccctc 1620
aaactggcag atgcacggtt acgatgcgcc catctacacc aacgtgacct atcccattac 1680
ggtcaatccg ccgtttgttc ccacggagaa tccgacgggt tgttactcgc tcacatttaa 1740
tgttgatgaa agctggctac aggaaggcca gacgcgaatt attttgatg gcgttaactc 1800
ggcgtttcat ctgtggtgca cgggcgctg ggtcggttac ggccaggaca gtcgtttgcc 1860
gtctgaattt gacctgagcg catttttacg cgccggagaa aaccgcctcg cggtgatggt 1920
gctgcgctgg agtgacggca gttatctgga agatcaggat atgtggcgga tgagcggcat 1980
tttccgtgac gtctcgttgc tgcataaacc gactacacaa atcagcgatt ccatgttgc 2040
cactcgcttt aatgatgatt cagccgcgc tgtactggag ctgaagttc agatgtgcgg 2100
cgagttgcgt gactacctac gggtaacagt ttctttatgg caggtgaaa cgcaggtcgc 2160
cagcggcacc gcgcctttcg gcggtgaaat tatcgatgag cgtggtggtt atgccgatcg 2220
cgtcacacta cgtctgaacg tcgaaaaccc gaaactgtgg agcgccgaaa tcccgaatct 2280
ctatcgtgcg gtggttgaac tgcacaccgc cgacggcacg ctgattgaag cagaagcctg 2340
cgatgtcggt ttccgcgagg tgcggattga aaatggtctg ctgctgctga acggcaagcc 2400
gttgctgatt cgaggcgtta accgtcacga gcatcatcct ctgcatggtc aggtcatgga 2460
tgagcagacg atggtgcagg atatcctgct gatgaagcag aacaacttta cgccgtgcg 2520
ctgttcgcat tatccgaacc atccgctgtg gtacacgctg tgcgaccgct acggcctgta 2580
```

```
tgtggtggat gaagccaata ttgaaaccca cggcatggtg ccaatgaatc gtctgaccga 2640
tgatccgcgc tggctaccgg cgatgagcga acgcgtaacg cgaatggtgc agcgcgatcg 2700
taatcacccg agtgtgatca tctggtcgct ggggaatgaa tcaggccacg gcgctaatca 2760
cgacgcgctg tatcgctgga tcaaatctgt cgatccttcc cgcccggtgc agtatgaagg 2820
cggcggagcc gacaccacgg ccaccgatat tatttgcccg atgtacgcgc gcgtggatga 2880
agaccagccc ttcccggctg tgccgaaatg gtccatcaaa aaatggcttt cgctacctgg 2940
agagacgcgc ccgctgatcc tttgcgaata cgcccacgcg atgggtaaca gtcttggcgg 3000
tttcgctaaa tactggcagg cgtttcgtca gtatccccgt ttacagggcg gcttcgtctg 3060
ggactgggtg gatcagtcgc tgattaaata tgatgaaaac ggcaacccgt ggtcggctta 3120
cggcggtgat tttggcgata cgccgaacga tcgccagttc tgtatgaacg gtctggtctt 3180
tgccgaccgc acgccgcatc cagcgctgac ggaagcaaaa caccagcagc agttttttcca 3240
gttccgtttta tccgggcaaa ccatcgaagt gaccagcgaa tacctgttcc gtcatagcga 3300
taacgagctc ctgcactgga tggtggcgct ggatggtaag ccgctggcaa gcggtgaagt 3360
gcctctggat gtcgctccac aaggtaaaca gttgattgaa ctgcctgaac taccgcagcc 3420
ggagagcgcc gggcaactct ggctcacagt acgcgtagtg caaccgaacg cgaccgcatg 3480
gtcagaagcc gggcacatca gcgcctggca gcagtggcgt ctggcggaaa acctcagtgt 3540
gacgctcccc gccgcgtccc acgccatccc gcatctgacc accagcgaaa tggattttg 3600
catcgagctg ggtaataagc gttggcaatt taaccgccag tcaggctttc tttcacagat 3660
gtggattggc gataaaaaac aactgctgac gccgctgcgc gatcagttca cccgtgcacc 3720
gctggataac gacattggcg taagtgaagc gacccgcatt gaccctaacg cctgggtcga 3780
acgctggaag gcggcgggcc attaccaggc cgaagcagcg ttgttgcagt gcacggcaga 3840
tacacttgct gatgcggtgc tgattacgac cgctcacgcg tggcagcatc aggggaaaac 3900
cttatttatc agccggaaaa cctaccggat tgatggtagt ggtcaaatgg cgattaccgt 3960
tgatgttgaa gtggcgagcg atacaccgca tccggcgcgg attggcctga actgccagct 4020
ggcgcaggta gcagagcggg taaactggct cggattaggg ccgcaagaaa actatcccga 4080
ccgccttact gccgcctgtt ttgaccgctg ggatctgcca ttgtcagaca tgtataccccc 4140
gtacgtcttc ccgagcgaaa acggtctgcg ctgcgggacg cgcgaattga attatggccc 4200
acaccagtgg cgcggcgact tccagttcaa catcagccgc tacagtcaac agcaactgat 4260
ggaaaccagc catcgccatc tgctgcacgc ggaagaaggc acatggctga atatcgacgg 4320
tttccatatg gggattggtg gcgacgactc ctggagcccg tcagtatcgg cggaattcca 4380
gctgagcgcc ggtcgctacc attaccagtt ggtctggtgt caaaaaagat ctggaggtgg 4440
tggcagcagg ccttggcgcg ccggatcctt aattaacaat tgaccggtaa taataggtag 4500
ataagtgact gattagatgc attgatccct cgaccaattc cggttatttt ccaccatatt 4560
gccgtctttt ggcaatgtga gggcccggaa acctggccct gtcttcttga cgagcattcc 4620
tagggggtctt tcccctctcg ccaaaggaat gcaaggtctg ttgaatgtcg tgaaggaagc 4680
agttcctctg gaagcttctt gaagacaaac aacgtctgta gcgacccttt gcaggcagcg 4740
gaaccccccca cctggcgaca ggtgcctctg cggccaaaag ccacgtgtat aagatacacc 4800
tgcaaaggcg gcacaacccc agtgccacgt tgtgagttgg atagttgtgg aaagagtcaa 4860
atggctctcc tcaagcgtat tcaacaaggg gctgaaggat gcccagaagg tacCCCattg 4920
tatgggatct gatctggggc ctcggtgcac atgctttaca tgtgtttagt cgaggttaaa 4980
aaacgtctag gccccccgaa ccacggggac gtggttttcc tttgaaaaac acgatgataa 5040
taccatgatt gaacaagatg gattgcacgc aggttctccg gccgcttggg tggagaggct 5100
attcggctat gactgggcac aacagacaat cggctgctct gatgccgccg tgttccggct 5160
gtcagcgcag gggcgcccgg ttctttttgt caagaccgac ctgtccggtg ccctgaatga 5220
actgcaggac gaggcagcgc ggctatcgtg gctggccacg acgggcgttc cttgcgcagc 5280
tgtgctcgac gttgtcactg aagcgggaag ggactggctg ctattgggcg aagtgccggg 5340
gcaggatctc ctgtcatctc accttgctcc tgccgagaaa gtatccatca tggctgatgc 5400
aatgcggcgg ctgcatacgc ttgatccggc tacctgccca ttcgaccacc aagcgaaaca 5460
tcgcatcgag cgagcacgta ctcggatgga agccggtctt gtcgatcagg atgatctgga 5520
cgaagagcat caggggctcg cgccagccga actgttcgcc aggctcaagg cgcgcatgcc 5580
cgacggcgag gatctcgtcg tgacccatgg cgatgcctgc ttgccgaata tcatggtgga 5640
aaatggccgc ttttctggat tcatcgactg tggccggctg ggtgtggcgg accgctatca 5700
ggacatagcg ttggctaccc gtgatattgc tgaagagctt ggcggcgaat gggctgaccg 5760
cttcctcgtg ctttacggta tcgccgctcc cgattcgcag cgcatcgcct tctatcgcct 5820
tcttgacgag ttcttctgag cgggactctg gggttcgcat cgataaaata aagattttta 5880
tttagtctcc agaaaaaggg gggaatgaaa gacccCacct gtaggtttgg caagctagct 5940
taagtaacgc cattttgcaa ggcatggaaa aatacataac tgagaataga gaagttcaga 6000
tcaaggtcag gaacagatgg aacagctgaa tatgggccaa acagatatc tgtggtaagc 6060
agttcctgcc ccggctcagg gccaagaaca gatggaacag ctgaatatgg gccaaacagg 6120
atatctgtgg taagcagttc ctgccccggc tcagggccaa gaacagatgg tccccagatg 6180
cggtccagcc ctcagcagtt ctagagaac catcagatgt ttccagggtg ccccaaggac 6240
ctgaaatgac cctgtgcctt atttgaacta accaatcagt tcgcttctcg cttctgttcg 6300
```

```
cgcgcttctg ctccccgagc tcaataaaag agcccacaac ccctcactcg gggcgccagt 6360
cctccgattg actgagtcgc ccgggtaccc gtgtatccaa taaaccctct tgcagttgca 6420
tccgacttgt ggtctcgctg ttccttggga gggtctcctc tgagtgattg actacccgtc 6480
agcggggggtc tttcattcat gcagcatgta tcaaaattaa tttggttttt tttcttaagt 6540
atttacatta aatggccata gttgcattaa tgaatcggcc aacgcgcggg gagaggcggt 6600
ttgcgtattg gcgctcttcc gcttcctcgc tcactgactc gctgcgctcg tcgttcggc 6660
tgcggcgagc ggtatcagct cactcaaagg cggtaatacg gttatccaca gaatcagggg 6720
ataacgcagg aaagaacatg tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg 6780
ccgcgttgct ggcgtttttc cataggctcc gcccccctga cgagcatcac aaaaatcgac 6840
gctcaagtca gaggtggcga aacccgacag gactataaag ataccaggcg tttcccctg 6900
gaagctccct cgtgcgctct cctgttccga ccctgccgct taccggatac ctgtccgcct 6960
ttctcccttc gggaagcgtg gcgctttctc atagctcacg ctgtaggtat ctcagttcgg 7020
tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc ccccgttcag cccgaccgct 7080
gcgccttatc cggtaactat cgtcttgagt ccaacccggt aagacacgac ttatcgccac 7140
tggcagcagc cactggtaac aggattagca gagcgaggta tgtaggcggt gctacagagt 7200
tcttgaagtg gtggcctaac tacggctaca ctagaagaac agtatttggt atctgcgctc 7260
tgctgaagcc agttaccttc ggaaaaagag ttggtagctc ttgatccggc aaacaaacca 7320
ccgctggtag cggtggtttt tttgtttgca agcagcagat tacgcgcaga aaaaaaggat 7380
ctcaagaaga tcctttgatc ttttctacgg ggtctgacgc tcagtggaac gaaaactcac 7440
gttaagggat tttggtcatg agattatcaa aaaggatctt cacctagatc cttttgcggc 7500
cgcaaatcaa tctaaagtat atatgagtaa acttggtctg acagttacca atgcttaatc 7560
agtgaggcac ctatctcagc gatctgtcta tttcgttcat ccatagttgc ctgactcccc 7620
gtcgtgtaga taactacgat acgggagggc ttaccatctg gccccagtgc tgcaatgata 7680
ccgcgagacc cacgctcacc ggctccagat ttatcagcaa taaaccagcc agccggaagg 7740
gccgagcgca gaagtggtcc tgcaacttta tccgcctcca tccagtctat taattgttgc 7800
cgggaagcta gagtaagtag ttcgccagtt aatagtttgc gcaacgttgt tgccattgct 7860
acaggcatcg tggtgtcacg ctcgtcgttt ggtatggctt cattcagctc cggttcccaa 7920
cgatcaaggc gagttacatg atcccccatg ttgtgcaaaa aagcggttag ctccttcggt 7980
cctccgatcg ttgtcagaag taagttggcc gcagtgttat cactcatggt tatggcagca 8040
ctgcataatt ctcttactgt catgccatcc gtaagatgct tttctgtgac tggtgagtac 8100
tcaaccaagt cattctgaga atagtgtatg cggcgaccga gttgctcttg cccggcgtca 8160
atacgggata ataccgcgcc acatagcaga actttaaaag tgctcatcat tggaaaacgt 8220
tcttcggggc gaaaactctc aaggatctta ccgctgttga gatccagttc gatgtaaccc 8280
actcgtgcac ccaactgatc ttcagcatct tttactttca ccagcgtttc tgggtgagca 8340
aaaacaggaa ggcaaaatgc cgcaaaaaag ggaataaggg cgacacggaa atgttgaata 8400
ctcatactct ccttttttca atattattga agcatttatc agggttattg tctcatgagc 8460
ggatacatat ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg cacatttc   8518
```

<210> 4
<211> 8175
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence for pICAST OMC

<400> 4

```
ctgcagcctg aatatgggcc aaacaggata tctgtggtaa gcagttcctg ccccggctca 60
gggccaagaa cagatggaac agctgaatat gggccaaaca ggatatctgt ggtaagcagt 120
tcctgccccg gctcagggcc aagaacagat ggtccccaga tgcggtccag ccctcagcag 180
tttctagaga accatcagat gtttccaggg tgccccaagg acctgaaatg accctgtgcc 240
ttatttgaac taaccaatca gttcgcttct cgcttctgtt cgcgcgcttc tgctccccga 300
gctcaataaa agagcccaca acccctcact cggggcgcca gtcctccgat tgactgagtc 360
gcccgggtac ccgtgtatcc aataaaccct cttgcagttg catccgactt gtggtctcgc 420
tgttccttgg agggtctcc tctgagtgat tgactacccg tcagcggggg tctttcattt 480
gggggctcgt ccgggatcgg agacccctg cccagggacc accgacccac caccgggagg 540
caagctggcc agcaacttat ctgtgtctgt ccgattgtct agtgtctatg actgatttta 600
tgcgcctgcg tcggtactag ttagctaact agctctgtat ctggcggacc cgtggtggaa 660
ctgacgagtt ctgaacaccc ggccgcaacc ctgggagacg tcccaggac tttggggggcc 720
gtttttgtgg cccgacctga ggaagggagt cgatgtggaa tccgaccccg tcaggatatg 780
tggttctggt aggagacgag aacctaaaac agttcccgcc tccgtctgaa tttttgcttt 840
cggtttggaa ccgaagccgc gcgtcttgtc tgctgcagca tcgttctgtg ttgtctctgt 900
```

```
ctgactgtgt ttctgtattt gtctgaaaat tagggccaga ctgttaccac tcccttaagt 960
ttgaccttag gtaactggaa agatgtcgag cggctcgctc acaaccagtc ggtagatgtc 1020
aagaagagac gttgggttac cttctgctct gcagaatggc caacctttaa cgtcggatgg 1080
ccgcgagacg gcacctttaa ccgagacctc atcacccagg ttaagatcaa ggtcttttca 1140
cctggcccgc atggacaccc agaccaggtc ccctacatcg tgacctggga agccttggct 1200
tttgaccccc ctccctgggt caagcccttt gtacacccta agcctccgcc tcctcttcct 1260
ccatccgccc cgtctctccc ccttgaacct cctcgttcga ccccgcctcg atcctccctt 1320
tatccagccc tcactccttc tctaggcgcc ggccgctcta gcccattaat acgactcact 1380
ataggccgat tcgaatcagg ccttggcgcg ccggatcctt aattaagcgc aattgggagg 1440
tggcggtagc ctcgagatgg gcgtgattac ggattcactg gccgtcgttt tacaacgtcg 1500
tgactgggaa aaccctggcg ttacccaact taatcgcctt gcagcacatc cccctttcgc 1560
cagctggcgt aatagcgaag aggcccgcac cgatcgccct tcccaacagt tacgcagcct 1620
gaatggcgaa tggcgctttg cctggtttcc ggcaccagaa gcggtgccgg aaagctggct 1680
ggagtgcgat cttcctgagg ccgatactgt cgtcgtcccc tcaaactggc agatgcacgg 1740
ttacgatgcg cccatctaca ccaacgtgac ctatcccatt acggtcaatc cgccgtttgt 1800
tcccacggag aatccgacgg gttgttactc gctcacattt aatgttgatg aaagctggct 1860
acaggaaggc cagacgcgaa ttattttttga tggcgttaac tcggcgtttc atctgtggtg 1920
caacgggcgc tgggtcggtt acggccagga cagtcgtttg ccgtctgaat ttgacctgag 1980
cgcattttta cgcgccggag aaaaccgcct cgcggtgatg gtgctgcgct ggagtgacgg 2040
cagttatctg gaagatcagg atatgtggcg gatgagcggc attttccgtg acgtctcgtt 2100
gctgcataaa ccgactacac aaatcagcga tttccatgtt gccactcgct ttaatgatga 2160
tttcagccgc gctgtactgg aggctgaagt tcagatgtgc ggcgagttgc gtgactacct 2220
acgggtaaca gtttctttat ggcagggtga aacgcaggtc gccagcggca ccgcgccttt 2280
cggcggtgaa attatcgatg agcgtggtgg ttatgccgat cgcgtcacac tacgtctgaa 2340
cgtcgaaaac ccgaaactgt ggagcgccga aatcccgaat ctctatcgtg cggtggttga 2400
actgcacacc gccgacggca cgctgattga agcagaagcc tgcgatgtcg gtttccgcga 2460
ggtgcggatt gaaaatggtc tgctgctgct gaacggcaag ccgttgctga ttcgaggcgt 2520
taaccgtcac gagcatcatc ctctgcatgg tcaggtcatg gatgagcaga cgatggtgca 2580
ggatatcctg ctgatgaagc agaacaactt taacgccgtg cgctgttcgc attatccgaa 2640
ccatccgctg tggtacacgc tgtgcgaccg ctacggcctg tatgtggtgg atgaagccaa 2700
tattgaaacc cacggcatgg tgccaatgaa tcgtctgacc gatgatccgc gctggctacc 2760
ggcgatgagc gaacgcgtaa cgcgaatggt gcagcgcgat cgtaatcacc cgagtgtgat 2820
catctggtcg ctggggaatg aatcaggcca cggcgctaat cacgacgcgc tgtatcgctg 2880
gatcaaatct gtcgatcctt cccgcccggt gcagtatgaa ggcggcggag ccgacaccac 2940
ggccaccgat attatttgcc cgatgtacgc gcgcgtggat gaagaccagc ccttcccggc 3000
tgtgccgaaa tggtccatca aaaaatggct ttcgctacct ggagagacgc gcccgctgat 3060
cctttgcgaa tacgcccacg cgatgggtaa cagtcttggc ggtttcgcta aatactggca 3120
ggcgtttcgt cagtatcccc gtttacaggg cggcttcgtc tgggactggg tggatcagtc 3180
gctgattaaa tatgatgaaa acggcaaccc gtggtcggct tacggcggtg attttggcga 3240
tacgccgaac gatcgccagt tctgtatgaa cggtctggtc tttgccgacc gcacgccgca 3300
tccagcgctg acggaagcaa aacaccagca gcagttttttc cagttccgtt tatccgggca 3360
aaccatcgaa gtgaccagcg aatacctgtt ccgtcatagc gataacgagc tcctgcactg 3420
gatggtggcg ctggatggta gccgctggc aagcggtgaa gtgcctctgg atgtcgctcc 3480
acaaggtaaa cagttgattg aactgcctga actaccgcag ccggagagcg ccgggcaact 3540
ctggctcaca gtacgcgtag tgcaaccgaa cgcgaccgca tggtcagaag ccgggcacat 3600
cagcgcctgg cagcagtggc gtctggcgga aaacctcagt gtgacgctcc ccgccgcgtc 3660
ccacgccatc ccgcatctga ccaccagcga aatggatttt tgcatcgagc tgggtaataa 3720
gcgttggcaa tttaaccgcc agtcaggctt tctttcacag atgtggattg gcgataaaaa 3780
acaactgctg acgccgctgc gcgatcagtt cacccgtgtc gatagatctg aacagaaact 3840
catttccgaa gaagacctag tcgaccatca tcatcatcat caccggtaat aataggtaga 3900
taagtgactg attagatgca tttcgactag atccctcgac caattccggt tattttccac 3960
catattccgg tcttttggca atgtgagggc ccggaaacct ggccctgtct tcttgacgag 4020
cattcctagg ggtctttccc ctctcgccaa aggaatgcaa ggtctgttga atgtcgtgaa 4080
ggaagcagtt cctctggaag cttcttgaag acaaacaacg tctgtagcga cccctttgcag 4140
gcagcggaac cccccacctg cgacaggtg cctctgcggc caaaagccac gtgtataaga 4200
tacacctgca aaggcggcac aaccccagtg ccacgttgtg agttggatag ttgtggaaag 4260
agtcaaatgg ctctcctcaa gcgtattcaa caaggggctg aaggatgccc agaaggtacc 4320
ccattgtatg ggatctgatc tggggcctcg gtgcacatgc tttacatgtg tttagtcgag 4380
gttaaaaaac gtctaagccc cccgaaccac ggggacgtgg ttttcctttg aaaaacacga 4440
tgataatacc atgaaaaagc ctgaactcac cgcgacgtct gtcgagaagt ttctgatcga 4500
aaagttcgac agcgtctccg acctgatgca gctctcggag ggcgaagaat ctcgtgcttt 4560
cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg ccgatggttt 4620
```

```
ctacaaagat cgttatgttt atcggcactt tgcatcggcc gcgctcccga ttccggaagt 4680
gcttgacatt ggggaattta gcgagagcct gacctattgc atctcccgcc gtgcacaggg 4740
tgtcacgttg caagacctgc ctgaaaccga actgcccgct gttctgcagc cggtcgcgga 4800
ggccatggat gcgatcgctg cggccgatct tagccagacg agcgggttcg gcccattcgg 4860
accgcaagga atcggtcaat acactacatg gcgtgatttc atatgcgcga ttgctgatcc 4920
ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgtccg tcgcgcaggc 4980
tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccgtgcacc tcgtgcacgc 5040
ggatttcggc tccaacaatg tcctgacgga caatggccgc ataacagcgg tcattgactg 5100
gagcgaggcg atgttcgggg attcccaata cgaggtcgcc aacatcttct tctggaggcc 5160
gtggttggct tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc 5220
aggatcgccg cggctccggg cgtatatgct ccgcattggt cttgaccaac tctatcagag 5280
cttggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg acgcaatcgt 5340
ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg cggccgtctg 5400
gaccgatggc tgtgtagaag tactcgccga tagtggaaac cgacgcccca gcactcgtcc 5460
gagggcaaag gaatagagta gatgccgacc gggatctatc gataaaataa aagatttttat 5520
ttagtctcca gaaaaagggg ggaatgaaag accccacctg taggtttggc aagctagctt 5580
aagtaacgcc attttgcaag gcatggaaaa atacataact gagaatagag aagttcagat 5640
caaggtcagg aacagatgga acagctgaat atgggccaaa caggatatct gtggtaagca 5700
gttcctgccc cggctcaggg ccaagaacag atggaacagc tgaatatggg ccaaacagga 5760
tatctgtggt aagcagttcc tgccccggct cagggccaag aacagatggt ccccagatgc 5820
ggtccagccc tcagcagttt ctagagaacc atcagatgtt ccagggtgc cccaaggacc 5880
tgaaatgacc ctgtgcctta tttgaactaa ccaatcagtt cgcttctcgc ttctgttcgc 5940
gcgcttctgc tccccgagct caataaaaga gcccacaacc cctcactcgg ggcgccagtc 6000
ctccgattga ctgagtcgcc cgggtacccg tgtatccaat aaaccctctt gcagttgcat 6060
ccgacttgtg gtctcgctgt tccttgggag ggtctcctct gagtgattga ctacccgtca 6120
gcggggggtct ttcattcatg cagcatgtat caaaattaat ttggtttttt ttcttaagta 6180
tttacattaa atggccatag ttgcattaat gaatcggcca acgcgcgggg agaggcggtt 6240
tgcgtattgg cgctcttccg cttcctcgct cactgactcg ctgcgctcgg tcgttcggct 6300
gcggcgagcg gtatcagctc actcaaaggc ggtaatacgg ttatccacag aatcagggga 6360
taacgcagga aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc 6420
cgcgttgctg gcgtttttcc ataggctccg cccccctgac gagcatcaca aaaatcgacg 6480
ctcaagtcag aggtggcgaa acccgacagg actataaaga taccaggcgt ttccccctgg 6540
aagctccctc gtgcgctctc ctgttccgac cctgccgctt accggatacc tgtccgcctt 6600
tctcccttcg ggaagcgtgg cgctttctca tagctcacgc tgtaggtatc tcagttcggt 6660
gtaggtcgtt cgctccaagc tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg 6720
cgccttatcc ggtaactatc gtcttgagtc caacccggta agacacgact tatcgccact 6780
ggcagcagcc actggtaaca ggattagcag agcgaggtat gtaggcggtg ctacagagtt 6840
cttgaagtgg tggcctaact acggctacac tagaagaaca gtatttggta tctgcgctct 6900
gctgaagcca gttaccttcg gaaaaagagt tggtagctct tgatccggca aacaaaccac 6960
cgctggtagc ggtggttttt ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc 7020
tcaagaagat cctttgatct tttctacggg gtctgacgct cagtggaacg aaaactcacg 7080
ttaagggatt ttggtcatga gattatcaaa aaggatcttc acctagatcc ttttaaatta 7140
aaaatgaagt ttgcggccgc aaatcaatct aaagtatata tgagtaaact tggtctgaca 7200
gttaccaatg cttaatcagt gaggcaccta tctcagcgat ctgtctattt cgttcatcca 7260
tagttgcctg actccccgtc gtgtagataa ctacgatacg gagggcttac catctggcc 7320
ccagtgctgc aatgataccg cgagacccac gctcaccggc tccagattta tcagcaataa 7380
accagccagc cggaagggcc gagcgcagaa gtggtcctgc aactttatcc gcctccatcc 7440
agtctattaa ttgttgccgg gaagctagag taagtagttc gccagttaat agtttgcgca 7500
acgttgttgc cattgctaca ggcatcgtgg tgtcacgctc gtcgtttggt atggcttcat 7560
tcagctccgg ttcccaacga tcaaggcgag ttacatgatc ccccatgttg tgcaaaaaag 7620
cggttagctc cttcggtcct ccgatcgttg tcagaagtaa gttggccgca gtgttatcac 7680
tcatggttat ggcagcactg cataattctc ttactgtcat gccatccgta agatgctttt 7740
ctgtgactgg tgagtactca accaagtcat tctgagaata gtgtatgcgg cgaccgagtt 7800
gctcttgccc ggcgtcaata cgggataata ccgcgccaca tagcagaact ttaaaagtgc 7860
tcatcattgg aaaacgttct cggggcgaa aactctcaag gatcttaccg ctgttgagat 7920
ccagttcgat gtaacccact cgtgcaccca actgatcttc agcatctttt actttcacca 7980
gcgtttctgg gtgagcaaaa acaggaaggc aaaatgccgc aaaaaaggga ataagggcga 8040
cacggaaatg ttgaatactc atactcttcc tttttcaata ttattgaagc atttatcagg 8100
gttattgtct catgagcgga tacatatttg aatgtattta gaaaaataaa caaatagggg 8160
ttccgcgcac atttc                                                    8175
```

<210> 5
<211> 8175
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence for pICAST OMN

<400> 5
```
ctgcagcctg aatatgggcc aaacaggata tctgtggtaa gcagttcctg ccccggctca 60
gggccaagaa cagatggaac agctgaatat gggccaaaca ggatatctgt ggtaagcagt 120
tcctgccccg gctcagggcc aagaacagat ggtccccaga tgcggtccag ccctcagcag 180
tttctagaga accatcagat gtttccaggg tgccccaagg acctgaaatg accctgtgcc 240
ttatttgaac taaccaatca gttcgcttct cgcttctgtt cgcgcgcttc tgctccccga 300
gctcaataaa agagcccaca acccctcact cggggcgcca gtcctccgat tgactgagtc 360
gcccgggtac ccgtgtatcc aataaaccct cttgcagttg catccgactt gtggtctcgc 420
tgttccttgg gagggtctcc tctgagtgat tgactacccg tcagcggggg tctttcattt 480
gggggctcgt ccgggatcgg gagacccctg cccagggacc accgacccac caccgggagg 540
caagctggcc agcaacttat ctgtgtctgt ccgattgtct agtgtctatg actgatttta 600
tgcgcctgcg tcggtactag ttagctaact agctctgtat ctggcggacc cgtggtggaa 660
ctgacgagtt ctgaacaccc ggccgcaacc ctgggagacg tcccagggac tttggggggcc 720
gtttttgtgg cccgacctga ggaagggagt cgatgtggaa tccgaccccg tcaggatatg 780
tggttctggt aggagacgag aacctaaaac agttcccgcc tccgtctgaa tttttgcttt 840
cggtttggaa ccgaagccgc gcgtcttgtc tgctgcagca tcgttctgtg ttgtctctgt 900
ctgactgtgt ttctgtattt gtctgaaaat tagggccaga ctgttaccac tcccttaagt 960
ttgaccttag gtaactggaa agatgtcgag cggctcgctc acaaccagtc ggtagatgtc 1020
aagaagagac gttgggttac cttctgctct gcagaatggc caacctttaa cgtcggatgg 1080
ccgcgagacg gcacctttaa ccgagacctc atcacccagg ttaagatcaa ggtctttttca 1140
cctggcccgc atggacaccc agaccaggtc ccctacatcg tgacctggga agccttggct 1200
tttgaccccc ctccctgggt caagcccttt gtacaccta agcctccgcc tcctcttcct 1260
ccatccgccc cgtctctccc ccttgaacct cctcgttcga ccccgcctcg atcctccctt 1320
tatccagccc tcactccttc tctaggcgcc ggccgctcta gcccattaat acgactcact 1380
atagggcgat tcgaatcagg ccttggcgcg ccgatcctt aattaagcgc aattgggagg 1440
tggcggtagc ctcgagatgg gcgtgattac ggattcactg gccgtcgttt tacaacgtcg 1500
tgactgggaa aaccctggcg ttacccaact taatcgcctt gcagcacatc ccccttttcgc 1560
cagctggcgt aatagcgaag aggcccgcac cgatcgccct tcccaacagt tacgcagcct 1620
gaatggcgaa tggcgctttg cctggtttcc ggcaccagaa gcggtgccgg aaagctggct 1680
ggagtgcgat cttcctgagg ccgatactgt cgtcgtcccc tcaaactggc agatgcacgg 1740
ttacgatgcg cccatctaca ccaacgtgac ctatcccatt acggtcaatc cgccgtttgt 1800
tcccacggag aatccgacgg gttgttactc gctcacattt aatgttgatg aaagctggct 1860
acaggaaggc cagacgcgaa ttatttttga tggcgttaac tcggcgtttc atctgtggtg 1920
caacgggcgc tgggtcggtt acggccagga cagtcgtttg ccgtctgaat ttgacctgag 1980
cgcattttta cgcgccggag aaaaccgcct cgcggtgatg gtgctgcgct ggagtgacgg 2040
cagttatctg gaagatcagg atatgtggcg gatgagcggc attttccgtg acgtctcgtt 2100
gctgcataaa ccgactacac aaatcagcga tttccatgtt gccactcgct ttaatgatga 2160
tttcagccgc gctgtactgg aggctgaagt tcagatgtgc ggcgagttgc gtgactacct 2220
acgggtaaca gtttctttat ggcagggtga aacgcaggtc gccagcggca ccgcgccttt 2280
cggcggtgaa attatcgatg agcgtggtgg ttatgccgat cgcgtcacac tacgtctgaa 2340
cgtcgaaaac ccgaaactgt ggagcgccga atcccgaat ctctatcgtg cggtggttga 2400
actgcacacc gccgacggca cgctgattga agcagaagcc tgcgatgtcg gtttccgcga 2460
ggtgcggatt gaaaatggtc tgctgctgct gaacggcaag ccgttgctga ttcgaggcgt 2520
taaccgtcac gagcatcatc tctgcatgg tcaggtcatg gatgagcaga cgatggtgca 2580
ggatatcctg ctgatgaagc agaacaactt aacgccgtg cgctgttcgc attatccgaa 2640
ccatccgctg tggtacacgc tgtgcgaccg ctacggcctg tatgtggtgg atgaagccaa 2700
tattgaaacc cacggcatgg tgccaatgaa tcgtctgacc gatgatccgc gctggctacc 2760
ggcgatgagc gaacgcgtaa cgcgaatggt gcagcgcgat cgtaatcacc gagtgtgat 2820
catctggtcg ctggggaatg aatcaggcca cggcgctaat cacgacgcgc tgtatcgctg 2880
gatcaaatct gtcgatcctt cccgcccggt gcagtatgaa ggcggcggag ccgacaccac 2940
ggccaccgat attatttgcc cgatgtacgc gcgcgtggat gaagaccagc cttcccggc 3000
tgtgccgaaa tggtccatca aaaaaatggct ttcgctacct ggagagacgc gcccgctgat 3060
cctttgcgaa tacgcccacg cgatgggtaa cagtcttggc ggtttcgcta aatactggca 3120
ggcgtttcgt cagtatcccc gtttacaggg cggcttcgtc tgggactggg tggatcagtc 3180
```

```
gctgattaaa tatgatgaaa acggcaaccc gtggtcggct tacggcggtg attttggcga 3240
tacgccgaac gatcgccagt tctgtatgaa cggtctggtc tttgccgacc gcacgccgca 3300
tccagcgctg acggaagcaa aacaccagca gcagtttttc cagttccgtt tatccgggca 3360
aaccatcgaa gtgaccagcg aatacctgtt ccgtcatagc gataacgagc tcctgcactg 3420
gatggtggcg ctggatggta agccgctggc aagcggtgaa gtgcctctgg atgtcgctcc 3480
acaaggtaaa cagttgattg aactacctga actaccgcag ccggagagcg ccgggcaact 3540
ctggctcaca gtacgcgtag tgcaaccgaa cgcgaccgca tggtcagaag ccgggcacat 3600
cagcgcctgg cagcagtggc gtctggcgga aaacctcagt gtgacgctcc ccgccgcgtc 3660
ccacgccatc ccgcatctga ccaccagcga aatggatttt tgcatcgagc tgggtaataa 3720
gcgttggcaa tttaaccgcc agtcaggctt tctttcacag atgtggattg gcgataaaaa 3780
acaactgctg acgccgctgc gcgatcagtt cacccgtgtc gatagatctg aacagaaact 3840
catttccgaa gaagacctag tcgaccatca tcatcatcat caccggtaat aataggtaga 3900
taagtgactg attagatgca tttcgactag atccctcgac caattccggt tattttccac 3960
catattgccg tcttttggca atgtgagggc ccggaaacct ggccctgtct tcttgacgag 4020
cattcctagg ggtctttccc ctctcgccaa aggaatgcaa ggtctgttga atgtcgtgaa 4080
ggaagcagtt cctctggaag cttcttgaag acaaacaacg tctgtagcga ccctttgcag 4140
gcagcggaac ccccacctg gcgacaggtg cctctgcggc caaaagccac gtgtataaga 4200
tacacctgca aaggcggcac aaccccagtg ccacgttgtg agttggatag ttgtggaaag 4260
agtcaaatgg ctctcctcaa gcgtattcaa caaggggctg aaggatgccc agaaggtacc 4320
ccattgtatg ggatctgatc tggggcctcg gtgcacatgc tttacatgtg tttagtcgag 4380
gttaaaaaac gtctaggccc cccgaaccac ggggacgtgg ttttcctttg aaaaacacga 4440
tgataatacc atgaaaaagc ctgaactcac cgcgacgtct gtcgagaagt ttctgatcga 4500
aaagttcgac agcgtctccg acctgatgca gctctcggag ggcgaagaat ctcgtgcttt 4560
cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg ccgatggttt 4620
ctacaaagat cgttatgttt atcggcactt tgcatcggcc gcgctcccga ttccggaagt 4680
gcttgacatt ggggaattta gcgagagcct gacctattgc atctcccgcc gtgcacaggg 4740
tgtcacgttg caagacctgc ctgaaaccga actgcccgct gttctgcagc cggtcgcgga 4800
ggccatggat gcgatcgctg cggccgatct tagccagacg agcgggttcg gcccattcgg 4860
accgcaagga atcggtcaat acactacatg gcgtgatttc atatgcgcga ttgctgatcc 4920
ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgtccg tcgcgcaggc 4980
tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc tcgtgcacgc 5040
ggatttcggc tccaacaatg tcctgacgga caatggccgc ataacagcgg tcattgactg 5100
gagcgaggcg atgttcgggg attcccaata cgaggtcgcc aacatcttct tctggaggcc 5160
gtggttggct tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc 5220
aggatcgccg cggctccggg cgtatatgct ccgcattggt cttgaccaac tctatcagag 5280
cttggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg acgcaatcgt 5340
ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg cggccgtctg 5400
gaccgatggc tgtgtagaag tactcgccga tagtggaaac cgacgcccca gcactcgtcc 5460
gagggcaaag gaatagagta gatgccgacc gggatctatc gataaaataa aagattttat 5520
ttagtctcca gaaaaagggg ggaatgaaag accccacctg taggtttggc aagctagctt 5580
aagtaacgcc attttgcaag gcatggaaaa atacataact gagaatagag aagttcagat 5640
caaggtcagg aacagatgga acagctgaat atgggccaaa caggatatct gtggtaagca 5700
gttcctgccc cggctcaggg ccaagaacag atggaacagc tgaatatggg ccaaacagga 5760
tatctgtggt aagcagttcc tgccccggct cagggccaag aacagatggt ccccagatgc 5820
ggtccagccc tcagcagttt ctagagaacc atcagatgtt tccagggtgc cccaaggacc 5880
tgaaatgacc ctgtgcctta tttgaactaa ccaatcagtt cgcttctcgc ttctgttcgc 5940
gcgcttctgc tccccgagct caataaaaga gcccacaacc cctcactcgg ggcgccagtc 6000
ctccgattga ctgagtcgcc cgggtacccg tgtatccaat aaaccctctt gcagttgcat 6060
ccgacttgtg gtctcgctgt tccttgggag ggtctcctct gagtgattga ctacccgtca 6120
gcgggggtct ttcattcatg cagcatgtat caaaattaat ttggtttttt ttcttaagta 6180
tttacattaa atggccatag ttgcattaat gaatcggcca acgcgcgggg agaggcggtt 6240
tgcgtattgg cgctcttccg cttcctcgct cactgactcg ctgcgctcgg tcgttcggct 6300
gcggcgagcg gtatcagctc actcaaaggc ggtaatacgg ttatccacag aatcagggga 6360
taacgcagga aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc 6420
cgcgttgctg gcgtttttcc ataggctccg ccccctgac gagcatcaca aaaatcgacg 6480
ctcaagtcag aggtggcgaa accgacagg actataaga taccaggcgt ttccccctgg 6540
aagctccctc gtgcgctctc ctgttccgac cctgccgctt accggatacc tgtccgcctt 6600
tctcccttcg gaagcgtgg cgctttctca tagctcacgc tgtaggtatc tcagttcggt 6660
gtaggtcgtt cgctccaagc tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg 6720
cgccttatcc ggtaactatc gtcttgagtc caacccggta agacacgact tatcgccact 6780
ggcagcagcc actggtaaca ggattagcag agcgaggtat gtaggcggtg ctacagagtt 6840
cttgaagtgg tggcctaact acggctacac tagaagaaca gtatttggta tctgcgctct 6900
```

Page 15

```
gctgaagcca gttaccttcg gaaaaagagt tggtagctct tgatccggca aacaaaccac 6960
cgctggtagc ggtggttttt ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc 7020
tcaagaagat cctttgatct tttctacggg gtctgacgct cagtggaacg aaaactcacg 7080
ttaagggatt ttggtcatga gattatcaaa aaggatcttc acctagatcc ttttaaatta 7140
aaaatgaagt ttgcggccgc aaatcaatct aaagtatata tgagtaaact tggtctgaca 7200
gttaccaatg cttaatcagt gaggcaccta tctcagcgat ctgtctattt cgttcatcca 7260
tagttgcctg actccccgtc gtgtagataa ctacgatacg ggagggctta ccatctggcc 7320
ccagtgctgc aatgataccg cgagacccac gctcaccggc tccagattta tcagcaataa 7380
accagccagc cggaagggcc gagcgcagaa gtggtcctgc aactttatcc gcctccatcc 7440
agtctattaa ttgttgccgg gaagctagag taagtagttc gccagttaat agtttgcgca 7500
acgttgttgc cattgctaca ggcatcgtgg tgtcacgctc gtcgtttggt atggcttcat 7560
tcagctccgg ttcccaacga tcaaggcgag ttacatgatc ccccatgttg tgcaaaaaag 7620
cggttagctc cttcggtcct ccgatcgttg tcagaagtaa gttggccgca gtgttatcac 7680
tcatggttat ggcagcactg cataattctc ttactgtcat gccatccgta agatgctttt 7740
ctgtgactgg tgagtactca accaagtcat tctgagaata gtgtatgcgg cgaccgagtt 7800
gctcttgccc ggcgtcaata cgggataata ccgcgccaca tagcagaact ttaaaagtgc 7860
tcatcattgg aaaacgttct tcggggcgaa aactctcaag gatcttaccg ctgttgagat 7920
ccagttcgat gtaacccact cgtgcaccca actgatcttc agcatctttt actttcacca 7980
gcgtttctgg gtgagcaaaa acaggaaggc aaaatgccgc aaaaaaggga ataagggcga 8040
cacggaaatg ttgaatactc atactcttcc tttttcaata ttattgaagc atttatcagg 8100
gttattgtct catgagcgga tacatatttg aatgtattta gaaaaataaa caaataggggg 8160
ttccgcgcac atttc                                                   8175
```

<210> 6
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> GS Linker

<400> 6
Gly Gly Gly Gly Ser
 1   5